Face Recognition

Face verification v.s. face recognition

Verification:

- Input image, name/ID pairs
- Output whether the input image is that of the claimed person.

Recognition:

- Has a database of K persons
- Get an input image
- Output ID if the image is any of the K persons.

One shot Learning

* Learning from one example to recognize the person again.

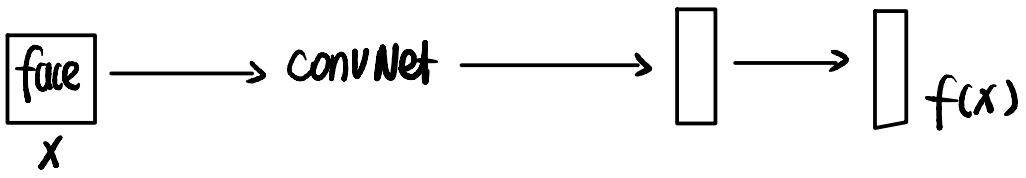
→ Learning a "similarity" function instead

$d(img1, img2)$ = degree of difference between images

if $d(img1, img2) \leq \tau$ ⟶ "same"

$> \tau$ ⟶ "different"

Siamese Network



Parameters of NN define an encoding $f(x^{(i)})$

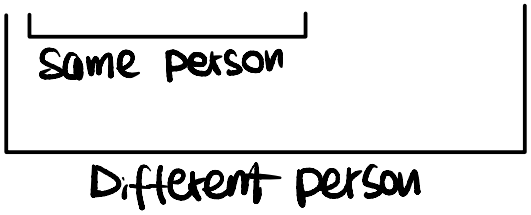
Learn parameters so that:

if $x^{(i)}, x^{(j)}$ are the same person, $\|f(x^{(i)}) - f(x^{(j)})\|_2^2$ is small

" different " " is large.

Triplet Loss

Triplets: Anchor img, positive img, negative img.



want: $\|f(A) - f(P)\|_2^2 + \alpha \leq \|f(A) - f(N)\|_2^2$

margin = this prevent nn from outputting same encoding for all images (trivial solution), which will always make the above expression valid.

Loss Function:

Given 3 images: A, P, N

$$L(A, P, N) = \max(\|f(A) - f(P)\|_2^2 - \|f(A) - f(N)\|_2^2 + \alpha, 0)$$

$$J = \sum_{i=1}^{m} L(A^{(i)}, P^{(i)}, N^{(i)})$$

Training set: e.g. 10k pictures of 1K persons.

Because of (P, A) pairs in triplets, you do need to collect multiple images of the same person.

Choosing the triplets A, P, N:

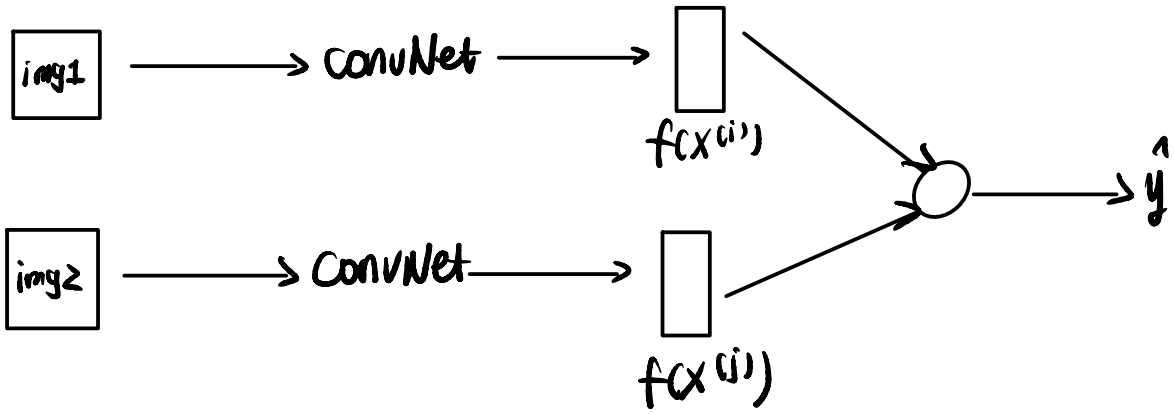
If A, P, N are chosen randomly, $d(A, P) + \alpha \leq d(A, N)$ is easily satisfied.

→ Choose triplets "hard" to train on

$$d(A, P) + \alpha \leq d(A, N)$$

$$\underbrace{d(A, P)}_{\downarrow} \approx \underbrace{d(A, N)}_{\uparrow}$$

Treat it AS Binary classification



$$\hat{y} = \sigma\left(\sum_{k=1}^{128} w_i \frac{|f(x^{(i)})_k - f(x^{(j)})_k| + b}{\frac{(f(x^{(i)})_k - f(x^{(j)})_k)^2}{f(x^{(i)})_k + f(x^{(j)})_k}}\right)$$

χ^2 Chi-Square.